

# Assessment for Learning: A Key to Motivation and Achievement

Rick Stiggins



**T**he role of schools has changed. Previously schools sorted students from the highest to the lowest achievers — that is, some succeeded at learning while others tumbled into chronic failure. But schools have evolved into places where all students are expected to meet increasingly rigorous academic standards.

This change is driven by the accelerating technical and ethnic evolution of our culture and the corresponding need for *all students* (not just those at the top of the achievement distribution) to become competent lifelong learners.

Our assessment practices historically have been designed to promote accountability by separating the successful from the unsuccessful learners and by highlighting their differences. However, given the new mission of ensuring universal competence, assessments now must support the learning of all students so that all can succeed at meeting standards. The result must be balanced assessment systems and a fundamental rethinking of the dynamics of assessment in effective schools.

Some important facets of this evolution in our priorities already have occurred. One example is our shift from almost total reliance on norm-referenced interpretation of test scores to far more frequent application of criterion-referenced interpretation of results. We have emerged from the era of comparing students with other students based on achievement to a time when we compare student performance to pre-set standards; and now we ask, who has and has not met standards?

Two other important shifts are just beginning to emerge. One is the desire to balance summative as-

sessments with formative assessments, and the second is to balance large-scale assessments with classroom assessments. Neither of these is a priority for mainstream assessments at this time. But they are becoming more important because of the recent discovery that profound achievement gains can be realized with effective, formative, classroom assessments.

While these developments are encouraging, there must be more profound changes. Perhaps the most fundamental of these changes must center on the way we judge the quality of an assessment. Historically, the challenge issued to the measurement community has been to produce dependable evidence of achievement. Thus attention traditionally has been lavished on the instruments (tests) that produce the scores. They had to be designed well and be built of high-quality ingredients to lead users to valid and reliable inferences about student learning. Decades of increasingly sophisticated technical advances have resulted in a deep understanding of how to produce, scale, and interpret test scores that consistently and accurately reflect our intended achievement targets. To be sure, this always must remain a foundation of quality assessment. Good data can underpin sound instructional decisions; bad data will lead to counterproductive decisions.





## This issue available in audio

Kappans can visit  
the PDK website —  
[www.pdkintl.org](http://www.pdkintl.org)  
— register, log in, and  
download the audio  
version of this article.

Use **Audio Edge**  
anytime you are  
on the go and want  
to fill your ears with  
the latest information.

All back issues of  
**Edge** also are avail-  
able online.

However, we must judge the quality of an assessment on far more than merely the dependability of its results. Quality also must be evaluated based on the effect of those results on the learner. The most dependable assessment in the world cannot be regarded as high quality if it has a counterproductive effect on learning or on students. For instance, an accurate score that causes a student to give up in hopelessness cannot be regarded as a quality assessment because it does more harm than good. Thus quality must become a function of the instrument and its score must be evaluated in terms of (or considered simultaneously with) the context and manner within which it is used. Quality control frameworks of the past have not taken this kind of effect on the learner into account.

Assessments must go beyond providing merely scores and corresponding judgments about student learning. Assessments must provide rich descriptions of the current state of student achievement. In other words, if assessments are to support improvements in student learning, their results must inform students how to do better the next time. This requires communicating results that transmit sufficient, understandable detail to guide the learner's actions. In such contexts, single scores or grades will not suffice.

In addition, if they are to support learning, assessments must evolve from being isolated events to becoming events that happen in an ongoing, interconnected series so that patterns in student learning will be revealed. In this way, both the learner and the teacher will be able to discern not only the student's current level of achievement, but also how much the student's capabilities have improved, which is a powerful booster for confidence and motivation.

Finally, to support learning, assessments must move beyond merely informing the instructional decisions of school leaders to informing decisions made by students and teachers, too. That means that we will need to design balanced assessment systems that serve diverse purposes by meeting the information needs of all decision makers. Historically, they have not done this.

This issue of *Edge* describes a vision of the future of assessment that accounts for each of these ingredients and advocates for bold movement into a new assessment future, revealing what these new assessments will do for student confidence, motivation, and achievement, as well as school effectiveness.



## Balanced Assessment Systems Serve Balanced Purposes

If schools are to be effective, educators must use the evidence gathered through assessments for two purposes: to inform instructional decisions and to encourage students to try to learn. In order to meet these purposes, assessment systems must yield accurate information about student learning for use at several levels of decision making, and they must be used in a manner that effectively manages the emotional dynamics of the assessment experience from the learner's point of view.

Managing the emotional dynamics of assessment means that educators must strive to create assessments that lead to productive reactions from both students and their teachers. For the student, a productive reaction leaves them confident and willing to keep trying. A counterproductive response has the student confused, frustrated, and ready to give up in hopelessness. For the teacher, the assessment is helpful if it reveals what comes next in the learning. The assessment is counterproductive when it leaves teachers with no idea of what to do next.

Productive assessment systems must serve many users. Decision makers at all levels need access to a variety of different kinds of information in different forms at different times. If any users' information needs are ignored or if the decision maker is provided with misinformation from inept assessments, ineffective decisions will result that will harm student confidence, motivation, and learning, as well as teacher efficacy.

For this reason, the first step in creating a quality assessment must be a clear sense of the information needs of the users. The assessor needs a clear sense of what kind of information is needed in order to know what kind of assessment must be conducted.

Table 1 analyzes the full range potential assessment users and uses within a school district. It begins by describing the assessment demands of the *classroom level*, where students, teachers, and parents make their instructional decisions. At the classroom level, assessment can be used both to support learning and to verify it. Next, the table progresses to the *program evaluation and support level*, where teacher leaders and teams, as well as principals, curriculum personnel, and others, rely on specific applications of assessment. In this case, one can identify students in need of help or evaluate program effects. And finally, we move to the *institu-*

## About the Author

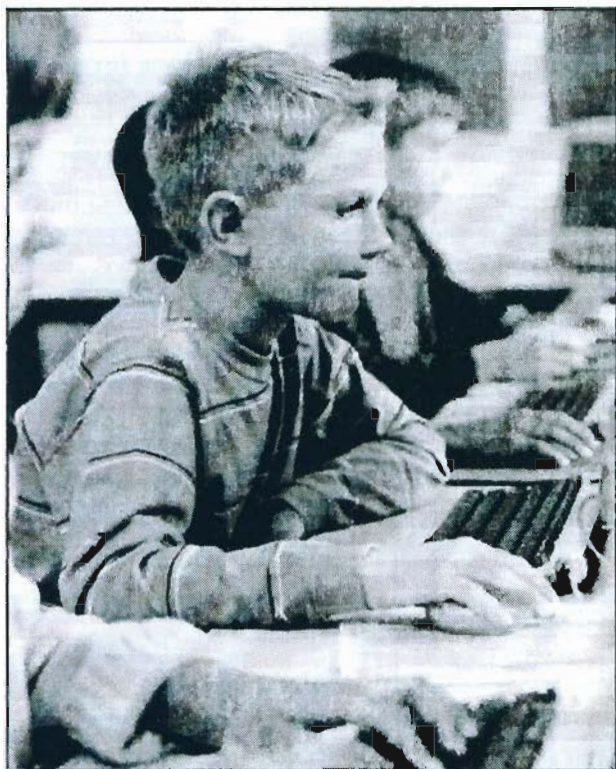
### Rick Stiggins

Rick Stiggins is founder and director of the ETS Assessment Training Institute in Portland, Oregon. ATI's flagship program, *Classroom Assessment FOR Student Learning*, provides a multi-media, collaborative, and hands-on way for teachers to become competent, confident, classroom assessors.

Stiggins has served on the faculties of Michigan State University, the University of Minnesota, and Lewis and Clark College. In addition, he directed test development at ACT in Iowa City and served as a senior researcher at the Northwest Regional Educational Laboratory.







**The most dependable  
assessment in the world  
cannot be regarded  
as high quality if it has  
a counterproductive  
effect on learning or  
on students.**

*tional and policy level* of assessment use, where resource allocation, program, policy, and other decisions are made by school, district, and community leaders. In this case, accountability decisions become most important.

To devise a truly useful assessment, one needs to know:

- What decision is to be made?
- By whom?
- What information will help them?

The point of table one is that the answers vary profoundly across the three levels. For instance, at the *classroom level*, the answers are as follows:

- Decision: What comes next in the learning?
- Made by: Students, teachers, parents.
- Needed information: Continuous evidence of each student's level of mastery of the steps leading to each standard.

The answers to the same three driving questions are different at the *level of program support*:

- Decision: Which students are meeting which standards?
- Made by: Teacher teams, teacher leaders, principals, curriculum personnel.
- Needed information: Periodic but frequent evidence of each student's mastery of each standard; comparable data permits aggregation across students.

It is critically important that we understand that both classroom and program levels of assessment and decision making are important, but they are different. One seeks to reflect progress toward mastery of standards, the other success in meeting standards. One involves only the adult decision makers, the other weaves students into the mix also. One focuses attention on the individual learner, the other on data summarized over learners. Productive assessment systems seek a synergy between the two by honoring the information needs of all important users.


And at the *institutional and policy level*:

- Decision: Are enough students meeting required standards?
- Made by: Superintendents, school boards, legislators, taxpayers.
- Needed information: Annual summaries of students' mastery of standards derived from accountability tests.

Thus the measurement community, as well as all


**Table 1: Comprehensive Analysis of  
Assessment Users and Uses.**

**Level 1: Classroom Assessment User: Student**


Important Question To be Answered	Information Needed to Answer the Question	Implications for the Assessment System
What am I supposed to learn?	Learning targets described in student-friendly language at the beginning of learning	Accurate assessments must reflect the learning targets students are given
What have I learned already, and what do I still need to work on?	Evidence must allow students to track progress and understand where they are now in relation to expectations at any point in time	Continuous sequence of accurate classroom assessments must provide descriptive feedback in student-friendly terms during learning
Have I met or am I progressing toward the important achievement standards?	Status regarding mastery of each standard in student-friendly language	Assessments must provide evidence of standards mastered periodically throughout the year
Have I met the state achievement expectations?	Status regarding meeting state standards in student-friendly language	Annual state assessments reporting standards mastered and not yet mastered
		



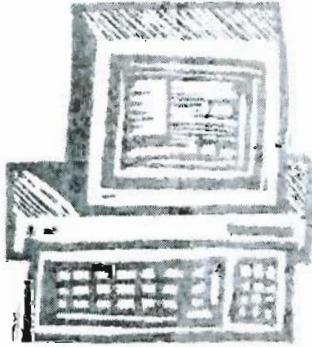
## Level 1: Classroom Assessment User: Teacher

Decisions to be Made	Information Needed	Implications for the Assessment System
What are my students supposed to learn?	Standards deconstructed into classroom targets leading, over time, to each standard; district curriculum maps of learning progression	All assessments must reflect these targets; it must be clear which target any assessment reflects
What have they learned already, and what do they still need to learn?	Continuous evidence revealing each student's current place in the learning progressions leading up to each standard	Continuous sequence of accurate classroom assessments used during the learning to provide picture of progress toward mastery of standards
Which students need special services?	Evidence of how students are doing in relation to grade- or age-level expectations	Assessments must provide evidence of students' relative status or progress to determine eligibility
Have my students met or are they progressing on the important achievement standards?	Status of each student's mastery of each standard	Periodic, interim benchmark assessments reflecting student mastery of standards throughout the year
Did they meet state achievement expectations?	Status regarding each student's mastery of each state standard	Annual assessments of each student's mastery of each state standard
		

## Level 1: Classroom Assessment User: Parent

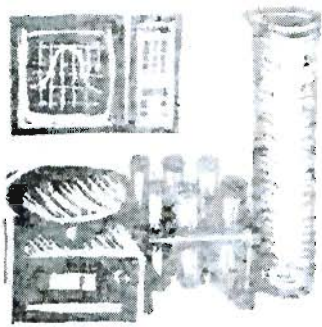
Decisions to be Made	Information Needed	Implications for the Assessment System
What is my child supposed to learn?	Learning targets in family-friendly language provided from the beginning of learning	Assessments must accurately reflect these targets
What has my child learned already, and what does she or he still need to learn?	Assessments providing information on current place in the progression to each learning target at any point in time	Continuous sequence of accurate classroom assessments used during the learning need to provide picture of progress
Is my child progressing satisfactorily in meeting the teacher's classroom learning expectations?	Information gained from my child's self-assessment, indications from the teacher or from my child	Periodic summative classroom assessments must feed into grade or summary of classroom standards met
Does my child need the services of a specialized program?	Student's learning in relation to grade- or age-level expectations	Assessment evidence needs interpretation in terms of expected achievement levels
Has my child met the state achievement expectations?	Status regarding meeting each state standard	Evidence of mastery of each standard gathered annually with accurate assessments
		

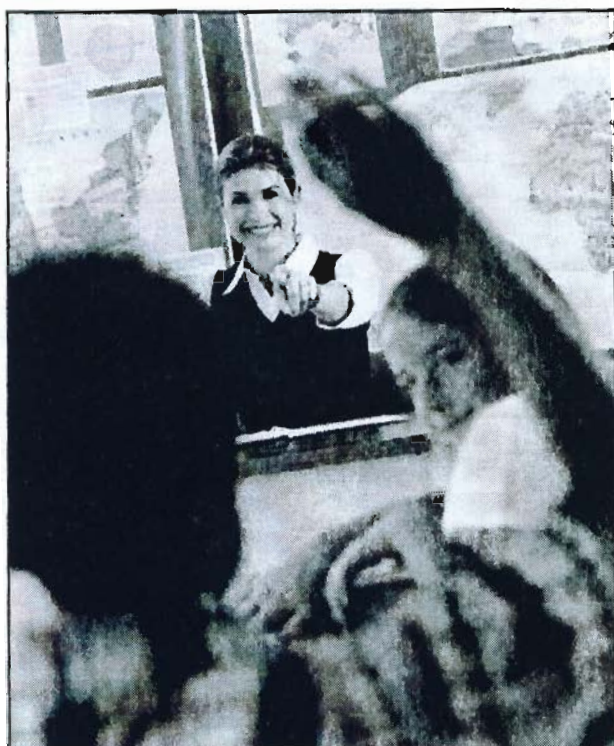
## Level 2: Instructional Support Users: Principal, Curriculum Leaders, Teacher Teams

Decisions to be Made	Information Needed	Assessment Implications
What standards are students expected to master by subject across our range of grade levels and classrooms?	Learning targets in the form of achievement standards organized by grade and subjects as they unfold within and across grade levels	Assessments must accurately reflect these standards and their associated classroom-level learning targets
Which of these standards are students mastering or progressing appropriately toward? Are there problem areas?	Information revealing patterns over time within the school year of achievement within and across teachers, grades, and subjects	Comparable evidence of student learning status collected periodically during the year
What standards are students to master across our classrooms, grades, and schools?	Standards mastered by grade and subject mapped within and across grade levels across schools	Assessments must accurately reflect these standards
Did enough of our students meet standards this year?	Proportion of students meeting each standard	Annual assessments reveal how each student does on each standard
		



### Level 3: Policy-Level Users: Superintendent, Various Policy Makers (School Boards, Legislators, Departments of Education, Business and Community Leaders)

Decisions	Information Needed	Implications for the Assessment System
What standards are to be met?	Learning targets in the form of achievement standards organized by grade and subject	Assessments must accurately reflect these standards
Which of these standards are students mastering or making appropriate progress toward in what schools?	Information revealing patterns of achievement within and across schools	Comparable evidence of student learning status collected periodically during the year
What standards are students expected to master in our schools?	Learning targets in the form of achievement standards organized by grade and subject	Assessments must accurately reflect these standards
How many of our students are meeting standards?	Scores reflecting patterns of achievement within and across schools and districts	Comparable evidence of student learning status collected periodically
Did enough of our students meet standards this year?	Proportion of students meeting each standard	Annual assessments show how each student scored on each standard
		



**... the starting place for  
developing a balanced  
assessment system is  
verifying the quality of  
the learning expectations  
on which it will rest.**

in school leadership and teaching positions, must see and understand the fundamental differences in the kinds of information needed across these levels of assessment use. No single assessment is capable of meeting the information needs of all of these various users. A productive, multi-level assessment system is needed to be sure that all instructional decisions are informed and made well. Table 1 breaks those needs into more specific person-by-person detail. Fail to meet the information needs of anyone on this list or fail to implement quality assessments at any level, and we place students directly in harm's way.

In other words, all parts of the system must make their unique contribution for schools to be effective. If assessment isn't working effectively in the classroom, program or policy levels of assessment can not pick up the slack. If bad decisions are being made during the learning, then there isn't an interim or annual assessment yet invented that can overcome the dire consequences for the learner. At the same time, equally important decisions are made at instructional support and policy levels that classroom assessment cannot replace.

The balanced assessment systems of the future — unlike the unbalanced systems of the past that were driven by standardized tests — can meet the information needs of all users.

We have available today all of the ingredients needed to use assessment productively. These ingredients include a large-scale assessment legacy, banks of high-quality test items for formative assessments, information management technologies, and the capacity to provide teachers with much needed professional development in classroom assessment. Now the challenge is to bring all of the parts to bear on behalf of student well-being.

### **Three Critical Foundations**

One structural foundation of any assessment system is the framework of achievement expectations to be reflected in the exercises and scoring schemes of its various component assessments. Whether those guiding achievement expectations are framed as state standards, local standards, a teacher's classroom standards, or the local curriculum designed to take students over time to those standards, certain keys to quality must be met. For example, the assessments must be:

- Centered on the truly important learnings of the field of study.



- Clearly and completely integrated into learning progressions within and across grades.
- Within developmental reach of the students who are to master them.
- Reflective of the best current thinking in the field.
- Manageable in number for mastery within the instructional context.
- Thoroughly mastered by those teachers charged with helping students master them.

If these criteria for sound standards are not met, then both high-quality assessment and effective instruction will remain beyond reach. Thus the starting place for developing a balanced assessment system is verifying the quality of the learning expectations on which it will rest. Until each local set of standards is in order, further consideration of assessment quality and use will be pointless.

The second foundation of an effective, balanced assessment is a commitment to developing and implementing standards-based schools. Faculty must understand what it means to design and offer standards-based instruction, and they must be committed to maximizing the success of each student in mastering the standards. Without these, focus will be missing, as will the willingness to invest in success.

The third foundation of a productive assessment system is accurate assessment. To yield accurate results, assessments must meet four standards of quality. They must 1) be designed to serve a specific, predetermined purpose (user and use); 2) arise from a specific, predetermined definition of achievement success; 3) be designed specifically to fit into each particular purpose and target context; and 4) communicate their results effectively (Stiggins et al. 2004). When an assessment is of high quality, it is sensitive enough to detect and accurately reflect changes in student achievement over time. When classroom, interim benchmark, or state assessments are not accurate or sensitive enough to track such changes, they will not contribute to productive assessment systems.

### A Revolution in Assessment for Student Motivation

When the mission of schools was merely to rank students (instead of also ensuring the competence of all students), the amount of time available to learn

was fixed: one year per grade. The amount learned by the end of that time was free to vary: some of us learned a great deal, some very little. Able learners built on past success to grow rapidly. However, students who failed to master the early prerequisites within the allotted time failed to learn that which followed. After 13 years of cumulative treatment in this manner, we were spread along an achievement continuum that literally labeled each student's rank in class upon graduation.

The motivational dynamics of this process were clear. From the very earliest grades, some students rode winning streaks to the top. Right from the start, they scored high on assessments. The emotional effect of this was to help them come to believe themselves to be capable learners — they became increasingly confident in school. That gave them the emotional strength to risk striving for more success because, in their minds, success was within reach if they tried. Notice, by the way, that the trigger for their learning success was their interpretation of their own success on assessments.

But other students scored very low on tests right from the beginning. This caused them to doubt their own capabilities as learners. They began to lose confidence, which, in turn, deprived them of the emotional reserves to continue to risk trying. Chronic failure was hard to hide and became embarrassing. It was better not to try. As their motivation waned, of course, achievement followed. Notice again how the learners' own interpretation of assessment results influenced their confidence and willingness to strive.

In these schools, if some students worked hard and learned a great deal, that was a positive result, as they would finish high in the rank order. And if some students gave up in the face of what they believed to be inevitable failure, that also was a necessary result, because they would occupy places very low in the rank order. The greater the spread of achievement from top to bottom, the more dependable would be the rank order. This is why, if a student gave up and stopped trying (even dropped out of school), it was regarded as that student's problem, not the teacher's or school's. The school's responsibility was to provide the opportunity to learn. If students didn't take advantage of the opportunity, that was not the system's responsibility.

The important lesson is that the student's emotional reactions to assessment results will determine what the student thinks, feels, and is motivated to do

in response to those results. They can respond in either of two motivational ways to any set of assessment results, one productive and the other not. The productive reaction has students saying, "I understand these results. I know what to do next to learn more. I can handle this. I choose to keep trying." The counterproductive response leaves students saying, "I don't know what these results mean for me. I have no idea what to do next. I can't handle this. I quit."

If society wants all students to meet standards, as specified above, then all students must believe they can meet those standards. They all must be confident enough about their chances of success to be motivated to take the risk of trying. Any other emotional state (such as the state of perpetual fear perpetrated in the schools of our youth) for any student is unacceptable. We can't have students who have yet to meet standards losing faith in themselves and giving up in futility.

As a result, assessment practices that permit, even encourage, some students to give up on learning must be replaced by those that engender hope and sustained effort for all students. In short, the entire emotional environment surrounding the experience of being evaluated must change for all, but especially for perennial low achievers. The driving emotional force of fear triggered by the prospect of an upcoming test now must be replaced by the emotions of optimism and persistence triggered by the belief that, "I can succeed at learning if I try." In other words, students must have continuous access to credible evidence of their own academic success.

I believe that school improvement experts have made the mistake of believing that the adults in the system are the most important assessment users and instructional decision makers. That is, they believed that schools become more effective as the adults make better instructional decisions. Certainly parents, teachers, school leaders, and policy makers make crucial decisions that can influence the quality of schools; and the more those decisions are based on data, the better. But, in fact, students may be even more important data-based instructional decision makers than are the adults.

Consider, for example, that students constantly are deciding if they can do the learning or not. They ask, can I get this, or is it just too hard for me? Is the learning worth the energy I must expend to attain it? Is the learning worth the risk of public failure? We must understand that, if students come down on the

wrong side of these crucial decisions and thus stop trying, it does not matter what the adults around them decide. In effect, students can render their teachers' instructional decisions null and void. They have it within their power to make the adults ineffective and to prevent them from doing anything about it. If a student decides that the learning is beyond reach for her or him or that the risk of public failure is too great and too embarrassing, then regardless of what adults do, there will be no learning.

So the essential issue for adults is, What can we do to help students answer the above questions in ways that keep them trying? We know how to do this, and it is not by intensifying the intimidation! Furthermore, we know what will happen to student achievement when we put effective classroom assessment practices in place. This leads to a key feature of the vision of excellence in assessment of the future.

### Productive Motivational Dynamics

Classroom assessment *for* student learning turns the classroom assessment process and its results into an instructional intervention designed to increase, not merely monitor, student learning, confidence, and motivation. Research evidence gathered in hundreds of studies conducted around the world over the past decade shows that the consistent application of principles of "assessment *for* learning" can give rise to unprecedented gains in student achievement, especially for perennial low achievers. The implications of such gains for raising test scores and closing achievement score gaps are profound.

One unique feature of the formative assessment *for* learning process is that it acknowledges the critical importance of the instructional decisions made by students and their teachers working as a team — it provides the information they need when they need it. In that context, students also become consumers of assessment information, using evidence of their own progress to understand what comes next for them.

Another important feature is its reliance on repeated self-assessments, each of which instructs the learner on how to improve performance on the next one. This kind of continuous descriptive feedback is provided strategically in amounts that students can address effectively, that is, in amounts that do not overwhelm students. This feedback builds progressively over time and thus helps students continue to believe that success is within reach if they keep trying.



Still another unique feature is its reliance on carefully drawn learning progressions or curriculum maps written in teacher-, student-, and family-friendly versions so that the trajectory (that is, what has been learned and what comes next) is clear to all throughout the learning. This, like the descriptive feedback described above, leads directly to our second reason for assessing: If we assess to motivate students to try, assessment *for* learning enables students by helping them watch themselves grow — by causing them to believe that success is within reach if they keep trying.

Thus the student's role in an assessment *for* learning environment is to strive to understand what success looks like and to use each assessment to determine how to do better the next time. Assessments become far more than merely one-time events attached to the end of the teaching. They become part of the learning process by keeping students posted on their progress and confident enough to continue striving. Students become partners in the self-assessment process by, for example, collaborating with their teachers in creating and using assessments like those they will be held accountable for later. This reveals to them the secrets to their own learning success while they are still learning. Students become partners in the accumulation of growth portfolios that reveal to them, their teachers, and their families the changes in their own achievement as it is happening. This builds confidence that ultimate success is within reach. Finally, students become partners in communicating about their own learning success as they rely on concrete evidence from their portfolios presented in student-led conferences to inform their families of their learning.

Evidence gathered around the world consistently reveals that, when such practices as these play out routinely in classrooms, effect sizes (test score gains) of as much as a full standard deviation can be realized, attributable to the application of formative classroom assessment *for* student learning.

In his original mastery learning research, Bloom and his students (1984) made extensive use of classroom assessment in support of learning in just the same terms as does the assessment *for* learning concept described here. Bloom reported subsequent gains in student test performance of one to two standard deviations. Black and Wiliam, in their 1998 watershed research review of more than 250 studies from around the world on the effect of classroom assessment, report gains of a half to a full standard



**... all parts of the system  
must make their  
unique contribution  
for schools to be  
effective.**





**The important lesson is that the student's emotional reactions to assessment results will determine what the student thinks, feels, and is motivated to do in response to those results.**

deviation, with the largest gains being realized by low achievers. Meisels and colleagues (2003) involved students in performance assessments and report gains of more than 1.5 standard deviations on subsequent tests. And finally, Rodriguez (2004) reports effects of similar size.

### **Assessment for Learning Illustrated**

Consider the following assessment experience from a high school English class. The assignment was to read three novels by the same author, develop a thesis statement, and write a term paper defending it, citing references from the literature.

To set students up for success, the teacher began by providing the students with a sample of an outstanding paper to read and analyze in order to determine what features made it outstanding. The next day in class they brainstormed and discussed what made it good, creating a written record of key features. Then the teacher gave them a sample of a very poor quality paper. Again they analyzed and evaluated its features in some detail. Comparing the two, they identified essential differences, ultimately collaborating in the development of a set of keys to quality that they transformed into a set of rating scales, each depicting the continuum of quality along a key dimension. Since students were partners in the process, the resulting scoring guide was written in student-friendly language and could be accompanied by examples of student work to illustrate each key.

Now the teacher was able to formulate instructional processes that dealt with each of the keys to success, one at a time and in sufficient depth to meet the learning needs of all students.

Furthermore, with the specific keys to quality clearly in mind, students were prepared to write the first drafts of their papers. They exchanged drafts, analyzing and evaluating each other's work and providing descriptive feedback on how to improve each. If students wanted feedback from their teacher on any particular dimension of quality, they could request and would receive it. Each revision yielded a paper of better quality. The paper was finished when the student knew it had reached an appropriate level of quality.

In the end, not every paper was outstanding. But most were of very high quality, and each student was confident of that fact before submitting their work for final evaluation and grading. If a student's



confidence as a writer had not been established by this time, this experience probably solidified it for many.

This teacher followed a proven set of principles of assessment *for learning* (Stiggins et al. 2004) that shows students what success looks like, how close they are coming to that target as they work, and how to continue to close the gap between their work and the agreed vision of excellence. Those principles are:

1. Start instruction by sharing a student-friendly version of the learning target(s) with the learners.
2. Accompany that with samples of student work that reveal the full range of quality.
3. Provide students with continuous access to descriptive feedback; help them see how to do better the next time.
4. Help them learn how to generate their own descriptive feedback.
5. Teach one facet of quality at a time.
6. Teach students the practice of focused revision — how to improve their work one facet at a time.
7. Teach students to understand, keep track of, and reflect on changes in their own proficiencies as they evolve over time.

When this kind of classroom assessment practice is used to inform and motivate students day to day in the classroom and is balanced with periodic standardized tests, the amount and quality of information brought to bear in the service of student success can be profound. Its effect on learning success is immediate and long lasting.

### Are We Ready for Balanced Assessment?

We know how to transform our assessment systems for use in standards-based schools. We understand what policy makers need to do in order to balance our assessment systems. We can continue to refine accountability oriented, large-scale assessments through sound research and development capacities, as we have for decades. In addition, with our banks of readily available assessment items and scoring capacities, along with instructional management technologies, we can help local districts develop interim, benchmark, or short-cycle assessments for use at the level of instructional program support. Finally, because there now are professional development programs in day-to-day classroom assessment, for the first time in the evolving history of

assessment in America, we can help local educators develop and use high-quality, day-to-day classroom assessments to support learning. So a totally balanced and integrated assessment system, with all parts working together in the service of student success, is within reach.

Not only can we help educators devise high-quality assessments, but we also know far more than ever before about how to use them effectively. The severe and chronic problem we can address now is the fact that very few teachers and almost no school administrators have been given the opportunity to learn about principles of sound assessment practice of any sort, let alone assessment *for learning*. While virtually all licensing standards require competence in assessment, typically neither preservice nor inservice teacher or administrator training programs include this kind of training (Crooks 1989; Stiggins 1999; Shepard et al. 2005). As a result of this lack of preparation:

- Educators are unable to differentiate among the various information needs of different assessment users, including students.
- Achievement targets remain written only at the level of state or district standards, rather than being translated into classroom-level learning progressions that lead up to each standard.
- The risk of inaccurate classroom assessments remains high.
- Feedback provided to students remains evaluative (such as grades), rather than helpfully descriptive.
- Students rarely are involved in self-assessment, tracking their own progress, or communicating their learning to others, all of which can give rise to profound learning gains.

Comprehensive, professional learning programs on sound assessment practice represent the only solution to this long-standing gap in the professional competence of our nation's teachers. These programs are needed for both preservice and inservice teachers. We must give new professionals the information needed to fulfill their increasingly complex assessment responsibilities, and we must provide older professionals who were not offered the opportunity to learn previously with those opportunities now. The essential competencies to be mastered are detailed in Table 2.

Preservice teacher preparation programs must provide foundational training in sound assessment

**Table 2: Indicators of Sound Classroom Assessment Practice\***

<p><b>1. Why Assess?</b></p> <p>Assessment Procedures and Results Serve Clear and Appropriate Purposes</p>	<ul style="list-style-type: none"> <li>a. Teachers understand who the users and uses of classroom assessment information are and know their information needs.</li> <li>b. Teachers understand the relationship between assessment and student motivation and craft assessment experiences to maximize motivation.</li> <li>c. Teachers use classroom assessment processes and results formatively (assessment for learning).</li> <li>d. Teachers use classroom assessment results summatively (assessment of learning) to inform someone beyond the classroom about students' achievement as of a particular point in time.</li> <li>e. Teachers have a comprehensive plan over time for integrating assessment for and of learning in the classroom.</li> </ul>
<p><b>2. Assess What?</b></p> <p>Assessments Reflect Clear and Valued Learning Targets</p>	<ul style="list-style-type: none"> <li>a. Teachers have clear learning targets for students; they know how to turn broad statements of content standards into classroom-level targets.</li> <li>b. Teachers understand the various types of learning targets they hold for students.</li> <li>c. Teachers select learning targets focused on the most important things students need to know and be able to do.</li> <li>d. Teachers have a comprehensive plan over time for assessing learning targets.</li> </ul>
<p><b>3. Assess How?</b></p> <p>Learning Targets Are Translated into Assessments that Yield Accurate Results</p>	<ul style="list-style-type: none"> <li>a. Teachers understand what the various assessment methods are.</li> <li>b. Teachers choose assessment methods that match intended learning targets.</li> <li>c. Teachers design assessments that serve intended purposes.</li> <li>d. Teachers sample learning appropriately in their assessments.</li> <li>e. Teachers write assessment questions of all types well.</li> <li>f. Teachers avoid sources of bias that distort results.</li> </ul>
<p><b>4. Communicate How?</b></p> <p>Assessment Results Are Managed Well and Communicated Effectively</p>	<ul style="list-style-type: none"> <li>a. Teachers record assessment of information accurately, keep it confidential, and appropriately combine and summarize it for reporting (including grades). Such summary accurately reflects current level of student learning.</li> <li>b. Teachers select the best reporting option (grades, narratives, portfolios, conferences) for each context (learning targets and users).</li> <li>c. Teachers interpret and use standardized test results correctly.</li> <li>d. Teachers effectively communicate assessment results to students.</li> <li>e. Teachers effectively communicate assessment results to a variety of audiences outside the classroom, including parents, colleagues, and other stakeholders.</li> </ul>
<p><b>5. Involve Students How?</b></p> <p>Students Are Involved in Their Own Assessment</p>	<ul style="list-style-type: none"> <li>a. Teachers make learning targets clear to students.</li> <li>b. Teachers involve students in assessing, tracking, and setting goals for their own learning.</li> <li>c. Teachers involve students in communicating about their own learning.</li> </ul>

\* Reproduced by permission from R. Stiggins, J. Arter, J. Chappuis, and S. Chappuis, *Classroom Assessment FOR Student Learning: Doing It Right — Using It Well* (Portland, Ore.: ETS Assessment Training Institute, 2004), p. 27.



practice. As preservice teachers come to understand the nature of the achievement targets their students need to master, so too must they learn how to use classroom assessment to track progress. This can be accomplished with sound coursework, but it would be even better if professors in teacher education courses would model sound assessment for learning practices in their own teaching.

Inservice programs for teachers must model the principles of sound professional development synthesized by the National Staff Development Council. Such learning experiences do the following:

- Focus on strategies proven to improve student learning (as effective, balanced assessment clearly has).
- Enable teachers to reflect on and improve their practice.
- Be an ongoing program, promoting continuous development over time.
- Rely on collaborative interaction and extensive hands-on practice.
- Merge into the normal work of teachers, helping them learn by working in their own, real, classroom context.
- Accommodate differences in teachers' starting places and rates of learning.

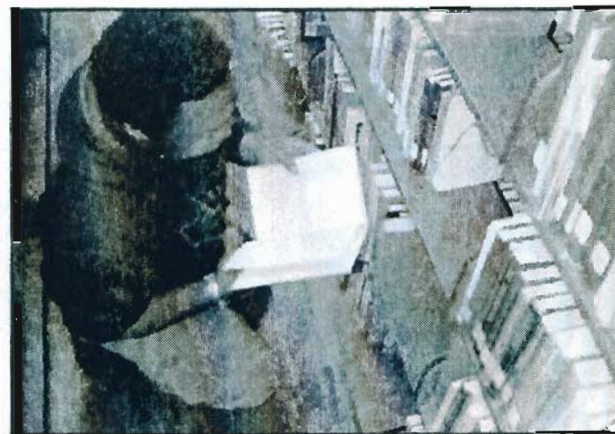
In our work at the ETS Assessment Training Institute, we have invested a decade and a half in creating and refining a learning team-based professional development program that lays this solid foundation of assessment literacy (Stiggins et al. 2004).

### Assessment at a Tipping Point

The current state of affairs is clear: We know what teachers need to know and understand to assess effectively day to day or year to year. We can provide them with the assessment tools and technologies needed to assess effectively.

It is clear what will happen to student learning when educators effectively use a balanced assessment program that motivates students. The achievement gains will be profound, especially for low achievers. And we know how to deliver the proper assessment competencies into the hands of all key users with efficient and effective professional development.

The only unanswered question is: Will practitioners be given the opportunity to learn to assess effectively? Historically, the answer has been a resound-



ing "No." As a result, the immense potential of assessment to support student learning has gone untapped. It need not be so. We have in hand a new vision of excellence in assessment that will tap the confidence, motivation, and learning potential that resides within every student.

### References

- Black, P., and William, D. "Assessment and Classroom Learning." *Educational Assessment: Principles, Policy and Practice* 5, no. 1 (1998): 7-74.
- Bloom, B. "The Search for Methods of Group Instruction as Effective as One-to-One Tutoring." *Educational Leadership* 41, no. 8 (1984): 4-17.
- Chappuis, S.; Stiggins, R.; Arter, J.; and Chappuis, J. *Assessment FOR Learning: An Action Guide for School Leaders*. 2nd ed. Portland, Ore.: ETS Assessment Training Institute, 2005.
- Crooks, T.J. "The Impact of Classroom Evaluation on Students." *Review of Educational Research* 58, no. 4 (1988): 438-481.
- Meisels, S.; Atkins-Burnett, S.; Xue, Y.; and Bickel, D.D. "Creating a System of Accountability: The Impact of Instructional Assessment on Elementary Children's Achievement Test Scores." *Educational Policy Analysis Archives* 11, no. 9 (2003). <http://cpaa.asa.edupaa/v11n9/>. Retrieved 21 April 2003.
- Rodriguez, M.C. "The Role of Classroom Assessment in Student Performance on TIMSS." *Applied Measurement in Education* 17, no. 1 (2004): 1-24.
- Shepard, L.; Hammerness, K.; Darling-Hammond, L.; Rust, F.; Snowden, J.B.; Gordon, E.; Gutierrez, C.; and Pacheco, J. In *Preparing Teachers for a Changing World: What Teachers Should Know and Be Able to Do*, edited by L. Darling-Hammond and J. Bransford. San Francisco: Jossey-Bass, 2005.
- Stiggins, R.J. "Evaluating Classroom Assessment Training in Teacher Education." *Educational Measurement: Issues and Practice* 18, no. 1 (1999): 23-27.
- Stiggins, R.; Arter, J.; Chappuis, J.; and Chappuis, S. *Classroom Assessment FOR Student Learning: Doing It Right — Using It Well*. Portland, Ore.: ETS Assessment Training Institute, 2004.



The Various Possible Links Between Achievement  
Targets and Assessment Methods

TARGET TO BE ASSESSED	ASSESSMENT METHOD			
	SELECTED RESPONSE	ESSAY	PERFORMANCE ASSESSMENT	PERSONAL COMMUNICATION
KNOWLEDGE MASTERY	Multiple choice, true/false, matching, and fill-in can sample mastery of elements of knowledge	Essay exercises can tap understanding of relationships among elements of knowledge	Not a good choice for this target--Three other options preferred	Can ask questions, evaluate answers and infer mastery-- but a time-consuming option
REASONING PROFICIENCY	Can assess understanding of basic patterns of reasoning	Written descriptions of complex problem solutions can provide a window into reasoning proficiency	Can watch students solve some problems and infer about reasoning proficiency	Can ask student to "think aloud" or can ask follow up questions to probe reasoning
SKILLS	Can assess mastery of the knowledge prerequisites to skillful performance--but cannot rely on these to tap the skill itself		Can observe and evaluate skills as they are being performed	Strong match when skill is oral communication proficiency; also can assess mastery of knowledge prerequisite to skillful performance
ABILITY TO CREATE PRODUCTS	Can assess mastery of knowledge prerequisite to the ability to create quality products--but cannot use these to assess the quality of products themselves		A strong match can assess; (a) proficiency in carrying out steps in product development, and (b) attributes of the product itself	Can probe procedural knowledge and knowledge of attributes of quality products--but not product quality
DISPOSITIONS	Selected response questionnaire items can tap student feelings	Open-ended questionnaire items can probe dispositions	Can infer dispositions from behavior and products	Can talk with students about their feelings